



Produto & Produção, vol. 11, n. 2, p. 85 - 103, jun. 2010

Confiabilidade em Questionários para Qualidade: Um Estudo com o Coeficiente Alfa de Cronbach

Henrique Rego Monteiro da Hora
Doutorando pela UFF
Professor do ISECENSA

Gina Torres Rego Monteiro
Professora Doutora da FioCruz
Dept. de Epidemiologia e Métodos Quantitativos em Saúde

Jose Arica
Professor Doutor da UENF
Laboratório de Engenharia de Produção

RESUMO

O setor de serviços ocupa na sociedade atual um papel de destaque em relação ao setor industrial, uma vez que os produtos industrializados tendem a possuir uma qualidade intrínseca semelhante, tornando o serviço oferecido junto com este produto um importante diferencial competitivo. O setor de serviço também se destaca em empresas que oferecem a sociedade serviços puros, ou seja, não agregados a nenhum outro produto. Dada a importância deste setor, é preciso avaliar se a qualidade dos serviços ofertados esta de acordo com as expectativas dos clientes. Este trabalho faz uma análise de confiabilidade de questionários para avaliar a qualidade de serviços em geral, por meio de um experimento realizado em um estudo de caso. Os resultados são significativos e levam a conclusão que o coeficiente utilizado deve ser mais difundido pela literatura nacional.

Palavras-chave: alfa de Cronbach, confiabilidade, qualidade em serviço.

1 INTRODUÇÃO

Bell (1973) constrói uma linha lógica pelo tempo e pela evolução tecnológica que caracteriza uma sociedade moderna ou pós-industrial (como o próprio autor denomina) baseado em dois parâmetros: o leque de serviços que ela oferece e o conforto que esses serviços possibilitam. No Quadro 1 é possível acompanhar a evolução das sociedades segundo Bell (1973).

	Sociedade		
	Pré-Industrial	Industrial	Pós-industrial
Característica	Extrativismo	Produção de Bens	Serviços
Energia	Força humana	Máquinas	Informação
Profissional	Sem qualificação	Semi-especializado	Habilitado
Limitações	Natureza	Natureza fabricada	Pessoas

Fonte: Baseado em Bell (1973).

Quadro 1: Principais características das sociedades.

É interessante notar a evolução das limitações de cada sociedade, onde há uma migração da natureza (chuva, estações do ano, fertilidade da terra) como componente crucial da produtividade, passando pelas máquinas e estruturas construídas pelo homem (natureza fabricada), chegando enfim a uma valorização do ser humano (*pessoa*), considerando-o peça importante na imagem e funcionamento da empresa.

A *Sociedade Pós-Industrial* prevista por Bell (1973) se aproxima grandemente da situação atual: de cada 100 trabalhadores americanos, 75 são empregados em organizações do setor de serviços, entretanto quando se considera apenas as indústrias desse setor, tem-se que 86% dos trabalhadores estão envolvidos com serviços, restando somente 14 em 100 para a fabricação de produtos (MIGUEL, 2001).

No Quadro 2 se ilustra a importância do setor de serviços pela quantidade de funcionários que as empresas alocam no setor administrativo. A GM, por exemplo, possui somente 22,5% da sua mão de obra destinada a produção direta de veículos e autopartes.

Empresa	Funcionários Administrativos
IBM	91,0 %
GM	77,5 %
Móbil	61,5 %
General Eletrics	60,0 %
Du Pont	57,1 %

Fonte: Baseado em Miguel (2001).

Quadro 2: Divisão da força de trabalho em algumas empresas.

Com o grande crescimento do setor, acirra-se a concorrência entre as organizações que prestam serviços. Admitindo-se que o nível de qualidade de uma organização é determinante para o seu sucesso e sua conseqüente sobrevivência no mercado, a medição da qualidade dos serviços prestados se torna uma importante tarefa para avaliar sua competitividade da mesma.

Os principais objetivos que levam uma organização a investir em qualidade são aumentar a fatia de mercado, conquistando novos clientes, e reduzir desperdícios (CARPINETTI et al, 2007). A lógica é que clientes satisfeitos representam faturamento, boa reputação, novos pedidos, melhores resultados para a empresa, garantia de emprego e remuneração para os funcionários. O contrário, clientes insatisfeitos, pode resultar em má reputação, dificuldade em conseguir novos pedidos, perda de faturamento e dificuldade em se manter no negócio.

Segundo dados do IBGE, o setor de serviços respondeu por 52% do PIB do Brasil em 2002 (MIGUEL & SALOMI, 2004; ROTONDARO & CARVALHO, 2007). A importância da preocupação com a qualidade do atendimento nesse setor de serviços pode ser comprovada pelos resultados de uma pesquisa que revelou um crescimento de 12% nos lucros das empresas que aplicaram programas para

melhoria da qualidade, contra 1% daquelas que não contaram com esse tipo de programa (LAS CASAS, 1992, apud MIGUEL, 2001). Outros estudos revelam que empresas que conseguiram diminuir a evasão de clientes em 5%, tiveram como consequência aumento do lucro entre 25% e 85% (MIGUEL, 2001).

Entretanto, qualidade em serviços é um conceito muito subjetivo e, por consequência, pouco preciso, pois a percepção individual de cada avaliador é proporcional às suas experiências anteriores, criando de certo modo uma expectativa sobre o serviço que ela receberá (PARASSURAMAN *et al*, 1985). Assim, quando se avalia o nível de satisfação do cliente, como medida de qualidade do serviço, deve-se solicitar dos mesmos seus pareceres em relação àquele (HAYES, 1995).

O presente trabalho propõe-se a uma análise de confiabilidade do questionário como instrumento de interrogação. O objetivo deste trabalho é ilustrar como estabelecer a confiabilidade de um questionário utilizado para a medição da qualidade em serviços por meio das avaliações de percepção, realizando um experimento a fim de aumentar a confiabilidade do questionário eliminando itens que de alguma forma, contribuem negativamente para o resultado final da análise de confiabilidade.

2 TEORIA CLÁSSICA DE MEDIÇÃO

Em toda medição o valor observado X é composto aditivamente por duas variáveis: o valor verdadeiro da medição “ V ”, e um erro aleatório de medição “ E ” (CARMINES & ZELLER, 1979; HAYES, 1995; ALLEN & YEN, 2003; CROCKER & ALGINA, 2006).

$$X = V + E \quad (1)$$

Analogamente, pode-se assumir que a variância total dos valores observados σ_X^2 está composta pela soma das variâncias dos valores verdadeiros σ_V^2 e a variância dos erros σ_E^2 (NICHOLS, 1999, HAYES, 1995).

$$\sigma_X^2 = \sigma_V^2 + \sigma_E^2 \quad (2)$$

2.1 Confiabilidade

Assim sendo, se a variância associada aos erros aleatórios diminui, o valor observado “ X ” se aproxima do valor verdadeiro “ V ”, o que representa maior precisão nas medições, conseqüentemente, maior confiabilidade no instrumento utilizado para a coleta dos dados (no caso, o questionário). Conceitualmente, a confiabilidade reflete o quanto os valores observados estão correlacionados aos verdadeiros valores (CARMINES & ZELLER, 1979; CROCKER & ALGINA, 2006).

A confiabilidade de um teste se define, equivalentemente, nos seguintes modos (NICHOLS, 1999):

- O quadrado da correlação entre os valores observados e os valores verdadeiros: (ρ_{xv}^2) ;

- A proporção da variância total dos valores observados em relação à variância dos valores reais: $(\sigma_x^2 / \sigma_v^2)$;
- Um menos a proporção da variância total dos valores observados pela variância do erro: $(1 - \sigma_x^2 / \sigma_E^2)$.

Entretanto, sendo σ_v^2 desconhecido, é impossível determinar a confiabilidade diretamente, calculando a correlação entre os valores observados e verdadeiros.

Por outro lado, Cronbach (1947) apresenta quatro definições para o termo confiabilidade e os respectivos métodos para sua estimação:

- *Definição 1 – coeficiente de estabilidade:* Confiabilidade é o grau com que o resultado de um teste permanece com diferenças inalteráveis individualmente em qualquer tratamento.
- *Definição 2 – coeficiente de estabilidade e equivalência:* Confiabilidade é o grau com que o resultado de um teste permanece com diferenças inalteráveis no universo e na amostra definida pelo teste.
- *Definição 3 – Coeficiente de equivalência:* Confiabilidade é o grau com que o resultado de um teste indica o status do indivíduo no universo e na amostra definida pelo teste.
- *Definição 4 – Auto-correlação hipotética:* Confiabilidade é o grau com que o resultado de um teste indica diferenças individuais em qualquer tratamento no presente momento.

Ao lançar estas definições Cronbach (1947) solicita que sejam discutidas e sugeridas novas e mais precisas definições. Aleen & Yen (2003), Hayes (1995) Crocker & Algina, (2006), e uma série de autores corroboram as definições de Cronbach (1947) e os métodos apresentados por ele para estimar a confiabilidade segundo cada definição:

- **Método do re-teste:** Aplicar o mesmo instrumento ao mesmo grupo. As aplicações são supostamente independentes, mas isso nem sempre pode ser feito. Mudanças nos resultados relativos são sempre tratadas como variância do erro, e não dos valores reais;
- **Método dos testes equivalentes:** Dois testes “equivalentes” ou “paralelos” devem ser aplicados com algum intervalo e sua correlação deve ser determinada. Qualquer mudança nos resultados deve ser tratada como variância do erro;
- **Método da divisão pela metade:** Supõe que o desvio padrão entre duas metades seja igual e que tenha alta correlação. Consiste em separar os resultados em duas metades e compará-las entre si.

Acrescenta-se a esta lista o método da consistência interna, onde são calculados estimadores para a confiabilidade.

2.2 Alfa de Cronbach

O coeficiente alfa de Cronbach foi apresentado por Lee J. Cronbach, em 1951, como uma forma de estimar a confiabilidade de um questionário aplicado em uma pesquisa. O alfa mede a correlação entre respostas em um questionário através da análise do perfil das respostas dadas pelos respondentes. Trata-se de uma correlação média entre perguntas. Dado que todos os itens de um questionário utilizam a mesma escala de medição, o coeficiente α é calculado a partir da variância dos itens individuais e da variância da soma dos itens de cada avaliador através da seguinte equação:

$$\alpha = \left(\frac{k}{k-1} \right) \times \left(1 - \frac{\sum_{i=1}^k s_i^2}{s_t^2} \right) \quad (3)$$

onde:

k corresponde ao número de itens do questionário;

s_i^2 corresponde a variância de cada item;

s_t^2 corresponde a variância total do questionário, determinada como a soma de todas as variâncias.

A Tabela 1 ilustra a aplicação passo a passo do coeficiente, onde cada coluna indica um item, cada linha indica um avaliador, e o encontro entre um item e um avaliador (X_{nk}) indica a resposta deste avaliador para este item, dentro da escala.

Tabela 1: Tabulação dos dados de questionário para cálculo do alfa de Cronbach.

Avaliadores	Itens						Total
	1	2	...	i	...	k	
1	X_{11}	X_{12}	...	X_{1i}	...	X_{1k}	X_1
2	X_{21}	X_{22}	...	X_{2i}	...	X_{2k}	X_2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
p	X_{p1}	X_{p2}	...	X_{pi}	...	X_{pk}	X_p
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	X_{n1}	X_{n2}	...	X_{ni}	...	X_{nk}	X_n
	s^2_1	s^2_2	...	s^2_n	...	s^2_k	S^2_t

Fonte: Adaptado de Cronbach, 2004.

Cada item, a priori, deve abordar uma única idéia de cada vez (GIL, 2002), isto é, os itens devem ser independentes. Se a resposta a determinado item se comporta de maneira parecida com a resposta de outro item, conclui-se que um explica o outro.

Assim sendo, para se ter um valor adimensional para representar a isenção dos erros aleatórios das observações (a confiabilidade), dividimos o estimador pela variabilidade total do questionário, chegando finalmente na fórmula apresentada por Cronbach em 1951 (CRONBACH, 2004).

A aplicação do alfa de Cronbach contempla alguns pressupostos, tais como:

- *O questionário deve estar dividido e agrupado em dimensões, ou seja, questões que tratam de um mesmo aspecto:* Parassuraman et al (1988), em seu clássico e seminal modelo de avaliação de qualidade em serviços, ServQual, após uma purificação de escala, agrupa 22 itens de

seu questionário em 5 dimensões (confiabilidade, presteza, garantia, empatia e tangibilidade);

- *O questionário deve ser aplicado a uma amostra significativa e heterogênea*: Quando se elabora um questionário para especialistas, a confiabilidade não pode ser aferida internamente, pois se subentende que avaliadores especialistas tendem a ter a mesma opinião sobre o assunto em discussão, diminuindo a variabilidade total do questionário e diminuindo assim o alfa. Neste caso, o método de confiabilidade entre avaliadores é mais recomendado do que a utilização do coeficiente alfa;
- *A escala já deve estar validada*: Há ainda na literatura científica uma confusão entre validade e confiabilidade. A validade está relacionada à verificação de se um instrumento realmente mede aquilo que se propõe a medir (o instrumento é válido para qual propósito?) Já confiabilidade está relacionada com a isenção de erros aleatórios, ou seja, em se repetir à verificação os resultados serão os mesmos?

Ainda não há um formalismo matemático para se dizer se uma escala é válida ou não, por esse motivo muitos pesquisadores avaliam a validade da escala pelo nível de confiabilidade desta. Exemplos deste procedimento podem ser encontrados em Freitas & Gonçalves (2004); Maçada *et al* (2000); Oviedo & Campo-Arias (2005), entre outros.

O conceito de “validade” trás subjetividade em si, pois ao afirmar que um instrumento de medição é válido, surge a pergunta: válido para qual propósito? A validade é o grau em que um determinado instrumento mede o que ele deveria medir (CARMINES & ZELLER, 1979). Validade é, em fim, o grau de acurácia ou exatidão do resultado de uma medição, i.e., é o quanto o resultado se aproxima do que se pretende medir (ALEEN & YEN, 2003).

Um exemplo claro da diferença entre os conceitos de validade e confiabilidade é ilustrado quando Cronin & Taylor (1992) questionam a validade do instrumento ServQual e propõe o ServPerf como sendo válido. Em vários trabalhos é possível verificar que em muitas medições da qualidade em serviços em diversos setores a escala ServQual é adaptada com um alto grau de confiabilidade, mas não tem a sua validade assegurada.

Salomi *et al* (2005) compararam os dois instrumentos (ServQual e ServPerf) utilizando o coeficiente alfa e verificaram um alto grau de confiabilidade para os dois instrumentos. Entretanto, um alto grau de confiabilidade não significa necessariamente que o instrumento seja válido.

2.2.1 Divulgação e utilização do coeficiente alfa

A grande utilização e aceitação no meio acadêmico do coeficiente alfa de Cronbach é um fator determinante para sua adoção como ferramenta para estimação da confiabilidade. O levantamento bibliométrico realizado por Peterson (1994) sobre a quantidade de publicações que utilizaram este coeficiente na literatura mundial é apresentado na Tabela 3. Não foram encontrados na academia dados consolidados mais atuais, e este trabalho não tem a ambição de realizar uma nova bibliometria.

Tabela 2: Quantidade de publicações sobre o alfa de Cronbach em fontes internacionais.

Fonte	Período	Quantidade
AMA/ACR Proceedings	1971-1992	113
Journal of Applied Psychology	1960-1992	670
Journal of Consumer Research	1974-1992	166
Journal of Marketing	1960-1992	238
Journal of Marketing Research	1964-1992	639
Journal of Personality and Social Psychology	1960-1992	724
Journal of Personality Assessment	1960-1992	586
Journal of the Academy of Marketing Science	1972-1992	387
Psychological Reports	1960-1992	418
Outras publicações	1970-1992	30
Manuscritos não publicados	1980-1992	315
Total		4.286

Fonte: Peterson, 1994.

Em consulta ao portal Scholar Google®, em junho de 2010, é possível verificar que mais de 12.000 citações já foram feitas ao artigo que lançou o coeficiente que ficou academicamente conhecido como alfa de Cronbach (1951).

No Brasil, seu uso ainda é pouco difundido nas revistas de Engenharia, mas alguns trabalhos são encontrados em periódicos da área de saúde, que frequentemente utiliza questionários em pesquisas (STREINER & NORMAN, 2008).

Tabela 3: Quantidade de publicações sobre o alfa de Cronbach em fontes nacionais.

Fonte	ISSN	Período	Quant.
Revista Produção on-line	1676-1907	2001-2008	1
Revista Produção	0103-6513	2002-2008	0
ENESEP		1996-2008	77
SIMPEP	1809-7189	2003-2007	1
Revista Gestão e Produção	0104-530X	2000-2008	0
Revista P&D em Engenharia de Produção	1679-5830	2003-2006	0
Revista de Administração de Empresas Eletrônica	1676-5648	2002-2007	0
Revista de Administração de Empresas	0034-7590	1961-2008	0
Outros periódicos	-	-	117
Total			196

Fonte: Elaboração Própria.

A Tabela 3 foi elaborada com os dados disponíveis em periódicos e bases científicas disponíveis para consulta on-line, e indexadas por, pelo menos, título do artigo.

É relevante ressaltar que, apesar da literatura científica a respeito das aplicações do coeficiente nas diversas áreas do conhecimento ser ampla e abrangente, ainda não existe um consenso entre os pesquisadores acerca da interpretação da confiabilidade de um questionário obtida a partir do valor deste coeficiente. Não há um valor mínimo definido para o coeficiente alfa de Cronbach ser aceito como bom, mas acha-se na literatura o valor de 0,70 como mínimo aceitável (FREITAS & CONGÇALVEZ, 2005; URDAN, 2001; OVIEDO & CAMPO-ARIAS, 2005; MILAN & TREZ, 2005). Existem, entretanto, autores que utilizam o alfa sem fazerem nenhuma menção a um mínimo (SALOMI *et al*, 2005; XEXEO, 2001; MIGUEL & SALOMI, 2004; MAÇADA *et al*, 2000; MATO & VEIGA, 2000).

2.2.2 Valores negativos para o coeficiente

O modelo para estimação alfa de Cronbach é válido para $\{\alpha \in \mathbb{R} \mid -\infty < \alpha \leq 1 \wedge \alpha \neq 0\}$, mas deve ser interpretado no intervalo entre 0 e 1, onde os valores negativos do alfa devem ser considerados como escalas sem confiança (ou seja, zero) (NICHOLS, 1999).

3 INSTRUMENTOS DE INTERROGAÇÃO

A aferição da satisfação do cliente pode se dar de várias formas, seja por pesquisa direta por meio de entrevistas, formulários e/ou questionários, seja por outros meios. Freitas (2005) levanta uma série de instrumentos para tal:

- **Formulários de Pesquisa (questionário):** para serem respondidos diretamente pelos clientes. Podem ser desenvolvidos baseados em métodos conhecidos, como ServQual, ServPerf, etc.;
- **Entrevistas:** realizadas pessoalmente ou por telefone;
- **Pesquisa de Mercado:** realizadas normalmente por um instituto de pesquisa independente;
- **Urna de Sugestões:** instrumento de comunicação diária com os clientes, capaz de captar elogios, críticas e sugestões;
- **Mesa Redonda com o Cliente:** reuniões planejadas entre a organização e clientes selecionados aleatoriamente;
- **Serviço de Atendimento ao Consumido (SAC):** linha telefônica ou e-mail disponibilizado para o cliente se manifestar.

Esses instrumentos de coleta de dados (ou instrumentos de medição) podem ser entendidos como os documentos através dos quais as perguntas são apresentadas aos respondentes, registrando-se as respostas e os dados obtidos (MATTAR, 2008). Nesse contexto, medição significa atribuir números a objetos ou eventos de acordo com regras. Ou ainda, a atribuição de uma escala numérica (quantitativa) ou descritiva (qualitativa) a um determinado evento ou objeto (CARMINES & ZELLER, 1979).

Gil (2008) apresenta três formas de levantamento de dados: a entrevista, o formulário, e o questionário. A entrevista se caracteriza por ser muito rica em informações e mas pouco poder de sintetização de dados. O questionário limita a liberdade de resposta, mas é muito fácil sumarizar dados com tratamento estatístico, e formulário esta entre essas duas realidades.

A principal característica do questionário, continua Gil (2008) é ser construído para ser preenchido sem auxílio de mão de obra especializada, e por isso, não necessitando de treinamento para realizar o levantamento de dados, e por isso, sendo a escolha da maioria das pesquisas por ser a forma mais barata de coletar informações.

Segundo Mattar (2008) o instrumento de coleta de dados é o documento através do qual as perguntas e questões são apresentadas aos respondentes e onde são registradas as respostas e dados obtidos. Todo o trabalho de planejamento e execução das etapas iniciais do processo de pesquisa se consolida

no instrumento de coleta de dados. O instrumento mais utilizado em pesquisas é o questionário.

Segundo Parasuraman *et al.* (1989), um questionário é tão somente um conjunto de questões, feito para gerar os dados necessários para se atingir os objetivos do projeto.

Mattar (2008) afirma que realizar medições é fundamental para a pesquisa de *marketing*. Os erros em pesquisas podem ser amostrais (com origem no processo amostral utilizado) ou não amostrais (com origem em diversos procedimentos errados ou equivocados da pesquisa).

Para a elaboração de um questionário é preciso seguir o seguinte roteiro:

- O problema e os objetivos da pesquisa;
- As hipóteses da pesquisa;
- A população a ser pesquisada;
- Os métodos de análise de dados escolhidos e disponíveis.

O autor supracitado ainda afirma que é preciso decidir o formato das respostas mais adequadas para avaliação, podendo ser:

- **Abertas:** Nas questões abertas, os respondentes ficam livres para responderem com suas próprias palavras e não se prendem a escolhas de alternativas. Exemplo: O que você pensa sobre o aborto?

Vantagens	Desvantagens
<ul style="list-style-type: none"> - Podem dar mais informações; - Muitas vezes mais ricas e detalhadas; - Podem dar uma informação inesperada. 	<ul style="list-style-type: none"> - Muitas vezes as respostas têm de ser interpretadas; - É preciso tempo para analisar as respostas; - Os respondentes muitas vezes têm preguiça de escrever.

Quadro 3: Vantagens e Desvantagens das Questões abertas. FONTE: Mattar (2008).

- **De múltipla escolha:** Nas questões de múltipla escolha, o respondente deve optar por uma das alternativas apresentadas no questionário. Exemplo: Marque a opção de 0 a 5 que achar de acordo com sua avaliação.

Vantagens	Desvantagens
<ul style="list-style-type: none"> - Facilidade de aplicação; - Rapidez de responder; - Possuem diversas alternativas. 	<ul style="list-style-type: none"> - Os respondentes podem ser influenciados pelas outras alternativas; - É preciso tempo para preparar as respostas corretas para não ocorrer um viés.

Quadro 4: Vantagens e desvantagens das questões de múltipla escolha. FONTE: Mattar (2008).

- **Dicotômicas:** Que apresenta duas opções, ou seja, caráter bipolar do tipo sim/não, concordo/não concordo, apoio/não apoio, normalmente uma terceira opção é oferecida, indicando desconhecimento ou falta de opinião.

Vantagens	Desvantagens
<ul style="list-style-type: none"> - Rapidez e facilidade de aplicação; - Apresentam pouca possibilidade de erro; - São objetivas. 	<ul style="list-style-type: none"> - Possibilidade de forçar a respostas pela opção apresentada; - São fortemente passíveis de erro dependendo da formulação das perguntas.

Quadro 5: Vantagens e desvantagens das questões dicotômicas. FONTE: Mattar (2008).

Algumas limitações de coleta de *feedback* dos clientes, reportam uma série de possíveis problemas decorrentes do uso do mesmo (CROSS *et al*, 1995):

- a amostragem pode não incluir usuários líderes, que são os primeiros a adotarem novidades;
- as perguntas podem condicionar os resultados;
- os clientes podem não se sentir à vontade para exprimir seus pontos de vista;
- numerosas estatísticas e gráficos, podem confundir as necessidades reais.

Embora a análise das técnicas conclua que o questionário constitui uma opção amplamente difundida como instrumento de coleta pelos inúmeros benefícios de seu uso (rápido de implementar, barato e sem exigência de treinamento de pessoal, entre outros), Hayes (1995) chama a atenção para a pouca literatura correspondente a como utilizar corretamente esse instrumento.

4 MATERIAIS E MÉTODOS

4.1 Classificação da pesquisa

Segundo Silva & Menezes (2001), a pesquisa é classificada quanto a sua natureza como aplicada, pois utiliza-se de técnicas já desenvolvidas para resolução de um problema local. Quanto à sua abordagem, a pesquisa é classificada como quantitativa, pois metrifica a qualidade de um serviço prestado e utiliza-se de técnicas estatísticas para chegar a conclusões. Quanto aos objetivos, a pesquisa é classificada como explicativa, pois procura aprofundar um conhecimento. Quanto aos procedimentos técnicos, a pesquisa é bibliográfica na fundamentação teórica e desenvolvimento da base conceitual do trabalho, e experimental na execução de seus procedimentos, realizando uma pesquisa documental no levantamento dos dados, que foram previamente coletados em outra pesquisa.

4.2 Procedimentos técnicos

A seguir, ilustra-se a verificação e purificação de escala de um questionário por meio da aplicação do coeficiente alfa de Cronbach em um questionário aplicado de acordo com a metodologia SoftMat, desenvolvida por Batista (2004).

Analisa-se a consistência interna do questionário da metodologia SoftMat desenvolvida e proposta por Batista (2004), tomando-se por base a avaliação do

software Régua e Compasso, Versão 2.41, disponibilizada em 10 de outubro de 2003.

A metodologia de avaliação SoftMat, disponível em <www.es.cefetcampos.br/softmat>, é de um instrumento de avaliação de *softwares* educacionais voltados para o ensino de matemática do ensino médio. Este é composto de um questionário, disposto em cinco blocos de questões. As questões levam em consideração aspectos técnicos das normas ISO (ISO/IEC 9126-1 e ISO/IEC 12119), aspectos educacionais em geral e aspectos particulares referentes à Matemática do Ensino Médio. Por meio desta metodologia são avaliados atributos de qualidade externa dos *softwares*.

A metodologia pode ser dividida em quatro fases distintas: 1. exposição do *software*; 2. resolução de atividades didáticas; 3. preenchimento do questionário; 4. cálculo do resultado.

O objetivo geral de todo o processo de avaliação é verificar a qualidade de um determinado *software* matemático quanto a sua aplicação em determinado conteúdo da grade do ensino médio. Entendendo-o como um produto adquirido, mas que tem algumas características de serviço (intangibilidade, simultaneidade, heterogeneidade (PARASURAMAN *et al*, 1985)), o uso do *software* pode ser comparado à prestação de um serviço.

4.2.1 Exposição do software

A fase de exposição do *software* tem como objetivo capacitar os avaliadores a usarem o *software* sozinhos, explorando e demonstrando suas funções e construindo um conhecimento com os alunos (avaliadores).

Esta fase é importante porque constitui um momento onde os conteúdos abordados podem – e devem – ser lembrados.

Dentro da metodologia SoftMat, procura-se dar condições plenas e necessárias para que o avaliador possa dar o seu parecer da forma mais consciente possível.

4.2.2 Resolução de atividades didáticas

Após a exposição do *software*, os avaliadores já se encontram em condições de resolverem uma série de exercícios comuns do dia-a-dia de uma sala de aula, e são instigados a fazerem um elo entre essas atividades didáticas e o plano de aula que eles estão acostumados a seguir, imaginando como poderia ser uma aula com o auxílio deste *software*.

Neste momento é desenvolvida uma atividade didática elaborada pela monitoria a fim de explorar as potencialidades do *software* em relação ao conteúdo matemático proposto.

A metodologia não prevê restrição quanto ao avaliador, mas no estudo de caso que está sendo apresentado, o grupo de avaliadores se compõe de licenciados e licenciados em matemática.

Dentro de um laboratório de informática, o(s) mediador(es) da oficina realizada para a avaliação do *software* em questão procura(m) construir o

conhecimento matemático proposto com auxílio de um *software* matemático. Nesta fase a interação entre monitores e avaliadores é muito bem vinda e a troca de experiências e habilidades, tanto com a matemática quanto com a informática são benéficas para a avaliação.

4.2.3 Preenchimento do questionário

Após a resolução das atividades didáticas, os avaliadores encontram-se familiarizados suficientemente com o *software* e com a proposta pedagógica para responder a um questionário que vai verificar a aderência do *software* com cada dimensão avaliada.

No presente estudo, a aplicação do questionário foi realizada na própria sala, de modo que o *software* ainda ficasse disponível para uso, mas a metodologia de Batista (2004) não prevê essa obrigatoriedade, podendo ser utilizada como algo particular, longe de uma oficina de aprendizado.

O questionário aplicado é composto por 5 dimensões, que Batista (2004) chama de blocos. Cada dimensão abrange os seguintes aspectos:

- *Bloco A*: questões relativas à documentação (documentação de descrição e manual do usuário, impresso ou *on line*);
- *Bloco B*: questões operacionais (relacionadas à instalação e utilização do *software*);
- *Bloco C*: questões relacionadas a características pedagógicas gerais (objetivos, usabilidade, conteúdos matemáticos e praticidade);
- *Bloco D*: questões relacionadas às propostas dos Parâmetros Curriculares do Ensino Médio (PCNEM) para Matemática;
- *Bloco E*: questões relativas à proposta pedagógica privilegiada no *software*.

A aplicação ainda em sala é estratégica e tem como objetivo disponibilizar o *software* ao avaliador para, no caso de ocorrer alguma dúvida quanto às suas potencialidades, limites ou aplicações, esta possa ser sanada de imediato com uma consulta à documentação do mesmo ou por uma nova investigação em suas funções.

O objetivo desse processo é preencher satisfatoriamente o questionário, usando ainda as atividades didáticas se necessário, bem como o próprio *software* ou o monitor/moderador.

4.2.4 Tratamento quantitativo dos dados

Após a coleta dos dados, aplicou-se a parte da metodologia para transformação das percepções dos avaliadores em índices de conformidade, tanto para cada bloco, como para o *software* completo.

A atribuição de valores dentro de uma escala nominal para uma escala numérica para as questões realizou-se da seguinte maneira (BATISTA, 2004):

- questões com resposta “Sim”, recebem valor 1;
- questões com resposta “Parcialmente com poucas restrições” recebem valor 0,75;

- questões com resposta “Parcialmente” recebem valor 0,5;
- questões com resposta “Parcialmente com muitas restrições” recebem valor 0,25;
- questões com resposta “Não”, recebem valor 0 (zero).

É importante ressaltar que em pesquisas, muitas vezes, os avaliadores não desejam, não sabem ou simplesmente não respondem a alguns itens do questionário. Nestas ocasiões, se usam pode-se usar alguns dos seguintes procedimentos: substituir as respostas em branco pelo valor zero; ignorar todas as demais respostas do avaliador, eliminando-o da análise; substituir as respostas em branco por um valor aleatório da escala de julgamento; ou, substituir as respostas em branco pela média dos valores respondidos no item (FREITAS & GONÇALVES, 2005).

Neste trabalho, são adotados os seguintes critérios para descarte i) se o número de questões respondidas por um avaliador for menor que 50%, todas as suas avaliações são eliminadas do cálculo; ii) se o número de questões respondidas por um avaliador for superior a 50%, seu julgamento em branco em um determinado item é substituído pela média dos valores dos outros avaliadores para o mesmo item. Esta forma de substituição, diversa da sugerida por Freitas & Gonçalves (2005), é preferível pois não altera as variâncias de cada item e contribui proporcionalmente à variância total do questionário. Esta medida assegura que os julgamentos de um determinado avaliador não sejam descartados indevidamente e que seu aproveitamento não seja prejudicial para a estimativa da confiabilidade.

4.5 Análise de confiabilidade e purificação da escala

Os dados coletados pelos questionários aplicados são tabulados em planilha eletrônica e as fórmulas modeladas do mesmo modo. Prefere-se a utilização de planilha eletrônica ao invés dos pacotes estatísticos profissionais disponíveis no mercado, que fazem cálculos similares, pelo motivo de acompanhamento do processo de cálculo e purificação da escala.

Parasuraman *et al.* (1989) citam que no desenvolvimento do ServQual as escalas foram purificadas, mas não definem precisamente o método utilizado para tal. Neste trabalho a purificação é realizada com o re-cálculo do coeficiente de confiabilidade, desconsiderando um único item, e observando o comportamento do mesmo. Ao final, sumariza-se a purificação eliminando todos os itens cuja eliminação contribui para o aumento da confiabilidade da medição.

5 RESULTADOS E DISCUSSÃO

Os blocos A, C, D e E obtiveram um bom nível de confiabilidade, considerado satisfatório e classificado na maioria das vezes como “Alto” segundo Freitas & Gonçalves (2005). Apenas o bloco B do questionário resultou um coeficiente muito abaixo do que se considera confiável (0,70) e classificado como “Muito Baixo”, segundo a mesma escala, chamando a atenção sobre os aspectos que possam ter levado ao baixo desempenho.

O Bloco B do questionário diz respeito aos quesitos técnicos do programa, a questões operacionais (relacionadas à instalação e utilização do *software*). Era esperado que licenciados e licenciandos não sejam os melhores avaliadores desses quesitos.

Já nos blocos restantes, vemos um bom desempenho no geral, em que o grupo de avaliadores pode realmente expressar sua opinião com excelência, pois o Bloco A diz respeito a documentação, e esta é preparada para qualquer usuário entender, os Blocos C, D e E dizem respeito a características pedagógicas, aos Parâmetros Curriculares Nacionais do ensino médio (PCNEM) e a proposta pedagógica do *software*, ou seja, áreas de conhecimento que profissionais da área de educação são altamente recomendados para darem sua opinião.

A intensidade da correlação entre os itens de um questionário pode ser verificada se esse coeficiente aumenta depois de eliminar um item da escala de medição (questionário). Se isso ocorrer, pode-se assumir que esse item não é altamente correlacionado com os outros itens da escala. Por outro lado, se o coeficiente diminuir pode ser assumido que esse item é altamente correlacionado com os outros itens da escala. Dessa forma, o alfa de Cronbach determina se a escala é realmente confiável, pois avalia como cada item reflete sua confiabilidade (SALOMI *et al*, 2005).

Neste trabalho, simula-se a eliminação de cada item de cada dimensão a fim de verificar o comportamento do coeficiente alfa e assim, determinar quais questões prejudicam a confiabilidade do questionário. Parassuraman *et al* (1988) chamam esta metodologia de purificação da escala.

Tabela 4: Resultado da purificação da escala.

Dimensão (Bloco)	Alfa inicial	Total de itens	Total de removidos	Remoção (%)	Alfa purificado	Incremento (%)
A	0,7380	18	5	27,78	0,8282	12,23
B	0,2949	11	9	81,82	0,6273	112,75
C	0,8046	27	7	25,93	0,8600	6,88
D	0,7324	7	3	42,86	0,8262	12,81
E	0,8168	8	1	12,50	0,8688	6,37

Fonte: Elaboração própria.

Na Tabela 4, pode-se observar que a quantidade de questões removidas para incremento do coeficiente de confiabilidade foi muito variável. No bloco A com a remoção de cinco itens (27,78% do total de itens do Bloco), o coeficiente passou de 0,7380 para 0,8282, ou seja, um incremento de 12,23% na confiabilidade antes da purificação.

O caso que mais nos chama a atenção é o Bloco B, onde o processo de purificação aponta que 81,82% dos itens deste bloco, se removidos, aumentariam em 112,75%, ou seja, resulta em um coeficiente com valor acima do dobro do coeficiente antes da purificação.

Este fenômeno que acontece com o Bloco B, de ter uma baixa confiabilidade, sugerindo a remoção de mais de 80% dos seus itens, reflete um alfa perto do razoável é justificado pela inabilidade apurada pelo perfil dos avaliadores em não serem aptos a avaliar questões operacionais relacionadas com o *software*, como “B.1. É fácil instalar o *software* corretamente?” ou “B.7. Na ocorrência de erros do sistema, o usuário tem fácil acesso às informações necessárias ao diagnóstico e solução do problema?”.

Outro tópico digno de atenção é o Bloco D que por possuir pequeno número de itens apresenta elevado percentual de remoção (42,86%), com um acréscimo na confiabilidade relativamente baixo (12,81%).

É importante notar que o processo de purificação de uma escala não é reduzido somente ao estudo de eliminação de itens. É necessário também uma rotação dos itens para identificar as dimensões (PARASSURAMAN *et al*, 1988), trabalho este que não pode ser realizado aqui pela pré-definição das dimensões como Blocos de questões propostas por Batista (2004).

Após a purificação de uma escala, esta deve ser revalidada em uma nova aplicação e o processo de avaliação de confiabilidade pela consistência interna se reinicia e se repete até o ponto em que se estabilize ou se chegue à confiabilidade desejada.

O questionário SoftMat, apesar de mostrar alguns blocos com um índice de confiabilidade baixo, ou então com uma purificação significativa, dificilmente poderá ter sua estrutura alterada, pois como instrumento para avaliação de *software*, necessariamente precisa abordar certos aspectos de acordo com as normas da ABNT, mesmo que os avaliadores não sejam os mais aptos a responderem um determinado conjunto de perguntas.

6 CONCLUSÃO

6.1 Quanto ao objetivo

Este artigo atinge seu objetivo ao ilustrar em detalhes a utilização do coeficiente alfa de Cronbach, realizando o experimento de eliminar itens de um questionário e observar o comportamento do mesmo coeficiente, o que é chamado de purificação de um questionário ou de uma escala (PARASSURAMAN *et al*, 1989).

Na área de engenharia, o alfa vem sendo utilizado cada vez mais, principalmente na área de engenharia da qualidade, como se verifica na bibliometria realizada no cenário nacional nos trabalhos consultados para esta pesquisa, para verificar a confiabilidade de medições. Trabalhos como este se fazem necessários para difusão e correta utilização do coeficiente.

Contudo, é importante ressaltar que esta técnica não substitui o papel do pesquisador em avaliar a pertinência de um item que avalia a percepção de um cliente acerca de algum aspecto que seja relevante. A análise de confiabilidade, bem como a purificação da escala devem auxiliar o pesquisador sobre a pertinência ou não de um item em determinado levantamento de dados.

6.2 Quanto ao trabalho realizado

Observaram-se ganhos significativos na confiabilidade das medições realizadas, com o descarte de um ou mais itens. Os resultados indicam que em um mesmo instrumento de interrogação, separado por dimensões da qualidade, é possível observar uma confiabilidade alta em uma parte, e uma extremamente baixa em outra, indicando ao pesquisador que ele deve avaliar a pertinência ou não de determinados itens.

O baixo desempenho do Bloco B, que diz respeito aos aspectos técnicos, é uma exigência da norma, e a permanência dos itens que representam um baixo desempenho na confiabilidade é necessária para contemplar todos os aspectos da qualidade em *software*.

6.3 Quanto aos trabalhos futuros

Parassuraman *et al* (1989), durante o seu processo de purificação da escala ServQual, utiliza um conceito de rotacionamento dos itens, onde procura reagrupar os itens de maneira que o coeficiente alfa torne-se maior. No trabalho de Parassuraman para avaliação da Qualidade em Serviços, começou-se com 10 dimensões para avaliação levantadas por Parassuraman *et al* (1985), mas após todo o procedimento, cinco dimensões foram removidas e a disposição dos itens foi completamente remodelada em comparação à proposta original.

Após este trabalho, é possível realizar um estudo de redefinição de dimensões e dos itens pertencentes a cada dimensão, rotacionando-os e reagrupando-os de forma a alcançar um maior índice de confiabilidade, mesmo sem a remoção de nenhum item.

Outro trabalho interessante a ser desenvolvido a partir deste é uma revisão de todos os itens cuja remoção foi sugerida, reaplicando os questionários em outras oficinas ou amostras pilotos, e por fim, apresentando um novo questionário mais confiável que o anterior.

ABSTRACT

The service sector occupies in nowadays's society a role highlighted in relation to the industrial sector since the industrial products tend to have a intrinsically quality similar, making the service offered with this product a significant competitive advantage. The service sector has also stood out in companies that offer pure services company, ie not aggregated with any other product. Given the importance of this sector, we must evaluate the services quality offered is in accordance with customer expectative. This paper has as aim a reliability analysis of questionnaires to assess the quality of services in general, by means of an experiment conducted on a case study. The results are significant and lead to the conclusion that the coefficient used should be encouraged by national literature.

Keywords: Cronbach's alfa, reliability, quality in service.

REFERÊNCIAS

ALLEN, M. J., YEN, W. M. *Introduction to measurement theory*. WaveLand Press, 2003.

BATISTA, S. C. F. *SOFTMAT: um repositório de ensino médio - um instrumento em prol de posturas mais conscientes na seleção de software*. Dissertação (Mestrado em Ciências de Engenharia). Campos dos Goytacazes, RJ, Universidade Estadual do Norte Fluminense – UEN, 229p. 2004.

BELL, D. *O Advento da sociedade pós-industrial*. Ed. Cultrix, 1973, São Paulo.

BESSION, J. L. *As Estatísticas: verdadeiras ou falsas?* In *Ilusão das Estatísticas*. Jean-Louis Besson (Org.). Editora UNESP. São Paulo, 1995.

CARMINES, E. G., ZELLER, R. A., *Reliability and validity assessment*. Sage University paper, 1979.

CARPINETTI, L. C. R., MIGUEL, P. A. C., GEROLAMO, M. C. *Gestão da qualidade ISO 9001:2000: princípios e requisitos*. Ed. Atlas, 110p. São Paulo, 2007.

CROCKER, L., ALGINA, J. *Introduction to classical & modern test theory*. Ed. Thonson, 571p, Florida 2006.

CRONBACH, J. L. *Coefficient alpha and the internal structure of tests*. V. 16. No. 3, pp. 297-334, *Psychometrika*, Setembro de 1951.

CRONBACH, J. L. *My current t procedures*. *Educational and Psychological Measurement*, Vol. 64 No. 3, Junho 2004.

CRONBACH, J. L. *Test "reliability": Its meaning and determination*. V. 12. No. 1, pp. 1 - 16, *Psychometrika*, Março de 1947.

CRONIN, J.; TAYLOR, S. *Measuring service quality: A reexamination and extension*. *Journal of Marketing*, v. 56, n. 3, p. 55-68, 1992.

CROSS, K. C., LYNCH, R. L., FEATHER, J. J. *A arte da reengenharia: o renascimento da empresa*. Jorge zahar editor, 307p. Rio de Janeiro, 1995.

FREITAS, A. L. P. *A qualidade de serviços no contexto da competitividade*. *Revista Produção on line*. Vol. 5, num. 1. UFSC – Florianópolis, SC, Março de 2005.

FREITAS, A. L. P., RODRIGUES, S. G. *A avaliação da confiabilidade de questionário: uma análise utilizando o coeficiente alfa de Cronbach*. XII SIMPEP – Bauru-SP, 2005.

GIL, A.C. *Como elaborar Projetos de Pesquisa*. Ed. Atlas. São Paulo - SP, 2002.

HAYES, B. E. *Medindo a satisfação do cliente: desenvolvimento e uso de questionários*. Quality Mark, 228p. Rio de Janeiro, 1995.

MAÇADA, A. C. G. *et al. Medindo a satisfação dos usuários de um sistema de apoio a decisão*. ENANPAD – Encontro Nacional de Administração, 2000.

MATO, C. A., VEIGA, R. T. *Avaliação da qualidade percebida de serviços: um estudo em uma organização não-governamental*. Caderno de Pesquisa em Administração, V. 07, N. 3, p. 27-42, São Paulo, Jul/Set de 2000.

MATTAR, F. N. *Pesquisa em marketing*. Ed. Atlas, 6. Ed. 347p. São Paulo, 2008.

MIGUEL, P. A. C. *Qualidade: enfoques e ferramentas*. Ed. Artliber, 263p. São Paulo, 2001.

MIGUEL, P. A. C., SALOMI, G. E. *Uma revisão dos modelos de medição de qualidade em serviços*. Revista Produção, V. 14. N. 1, p. 12-30. Rio de Janeiro, 2004.

MILAN, G. S., TREZ, G. *Pesquisa de satisfação: um modelo para planos de saúde*. RAE Eletrônica, Revista de Administração de Empresas. V. 4, N. 2, Art. 17, jul/dez 2005.

NICHOLS, D. P. *My Coefficient α is Negative*, SPSS Keywords, Number 68, 1999 disponível em <<http://www.ats.ucla.edu/STAT/SPSS/library/negalfa.htm>> acessado em 04/08/2006.

OVIEDO, H. C., CAMPO-ARIAS, A. *Aproximación al uso del coeficiente alfa de Cronbach*, Revista Colombiana de Psiquiatria. V. 34, N. 4. p. 572-580. 2005.

PARASSURAMAN V. ZEITHAMI, V.A. BERRY, L. L. *A conceptual model of service quality and its implications for future research*. Journal of Marketing, v. 49, n. 4, p. 41-50, 1985.

_____. *SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality*, Journal of Retailing, v. 64, n. 1, p. 12-40, 1989.

PETER, J. P., *Reliability: a review of psychometric basics and recent marketing practices*, Journal of Marketing Research, vol. 16, 6-17. 1979.

PETERSON, R. A. *A meta-analysis of Cronbach's coefficient alfa*, Journal of Consumer Research; ABI/INFORM Global, Set 1994

ROTONDARO, R. G., CARVALHO, M. M. *Qualidade em serviços* pp. 331-355 in: CARVALHO, M. M., PALADINI, E. P (org). *Gestão da qualidade: teoria e casos*. Ed. Campus, 355p. Rio de Janeiro, 2007.

SALOMI, G. G. E. *et al, SERVQUAL x SERVPERF: comparação entre instrumentos para avaliação da qualidade de serviços internos*. Gestão da Produção. V. 12 N. 2 p 279-293, maio-ago/2005.

SILVA, E. L. ; MENEZES, E. M. *Metodologia da pesquisa e elaboração de dissertação*. Florianópolis, 2001.

STREINER, D. L., NORMAN, G. R. *Health measurement scales: a practical guide to their development and use*. 4th edition, Oxford, 431p, 2008

TROCHIM, W. M. *The Research Methods Knowledge Base*, 2nd Edition. Internet WWW page, at URL: <<http://trochim.human.cornell.edu/kb/index.htm>> (version current as of August, 2005).

URDAN, A. T. *Qualidade de Serviços médicos na perspectiva do cliente*. Revista de Administração de Empresas, V. 41, N. 4 p. 44-55, out/dez 2001.

XEXEO, J. A. M., *Sistemas de Informação como Instrumento de Programa de Qualidade*, doutorado em Engenharia de Sistemas e Comunicação, COOPE, UFRJ, 2001.